Analyzing Experimental Data

Descriptive statistics (means, medians, sds, variance) vs. Inferential statistics

Could an observed difference between conditions/groups have occurred by chance?

The effect of error variance.

Need some objective way to determine the likelihood of an observed effect being due to chance: p-value.

Method for determining the significance of an observed difference

Hypothesis testing

Experimental hypothesis (there's a difference)

Null hypothesis (no difference, or no difference in the expected direction).

Hypothesis testing

Types of errors in hypothesis testing

Effect sizes

Hypothesis testing is a black-white distinction. You either are significant or not.

Usually want to also know the size of the difference: the EFFECT SIZE

Confidence interval of the difference between two groups

Hypothesis testing of the mean difference between TWO groups

t-test

Analogy – playing darts.

If you observe that the average score by player A is 100 and the average score by player B is 85, what percentage of the time would you expect player B to beat player A?

Need to know the consistency of each player's score, right???

Draw the two curves and determine what percentage of the time player B's values will be higher than player A's.

Goal is to determine whether the computed difference is significantly different from 0.

1. Calculate the means of the 2 groups

2. Calculate the s.e. of the mean difference

NOTE: The book is in error here!!!

See above for computations.

3. Find the calculated value of 't'

't' is an index of effect size.

Analogous to z-score but for small sample sizes

To find t, it's like finding z-scores...

4. Find the critical value of 't'

This depends on your alpha level - how different must they be to conclude 'statistically significant'?

Use a table (Appendix A2, p. 413).

Must know the df - proportional to sample size. Bigger samples means a better estimate.

$df = n_1+n_2 -2$

The sign of t

5. Compare the calculated t with the critical t.

Is $t_{calculated}$ more extreme (farther from 0) than $t_{critical}$.

Computational example:

gender and height from class

Within-subject analyses (paired-t test)

Only look at differences within a subject, not between.

Robustness of the t-test

The methods discussed assume

underlying distributions are normal

variance in each group is approximately equal

Deviations from these assumptions can invalidate your conclusions

Analysis of Variance (App. C).

The problem: inflated Type I error by using lots of t-tests.

2 levels, 1 t-test: alpha = .05

3 levels, 3 t-tests (AB AC BC), alpha = .14

4 levels, 6 t-tests (AB, AC, AD, BC, BD, CD), alpha = .26

ANOVA checks for the presence of any difference due to IV levels.

All levels checked simultaneously

Two outcomes: IV had no effect, or some level of the IV had an effect on the DV.

Hypothesis testing

H_0: $\mu_1=\mu_2=\mu_3=....=\mu_k$

H_a: not (H_0: $\mu_1=\mu_2=\mu_3=....=\mu_k$)

When you reject the null hypothesis, H_0, you do not know which of the means were different from which others

Post hoc tests required

Conceptually - how it works

Check if between-group variance is larger than within-group variance

similar to checking if total variance is larger than error variance

within-group variance is an estimate of error variance

between-group variance is error variance + a function of systematic variance

BOOK ERROR ABOVE

harken back to project 1

To test if the difference is significant, one uses an F-test

F = between group / within group

Note the effect of error variance.

REMINDER: We're dealing with inferential statistics here rather than the descriptive stats used in Project 1

Calculations

Sums of squares

$SStotal = \sum(x_i - GM)^2$

Note that total variance = SStotal / n-1

The total sum of squares = between group SS and within-group SS

Within group variance is an estimate of error variance

$SSwg = \sum (x_1 - \text{x-bar}_1)^2 + \sum(x_2 - \text{x-bar}_2)^2 + \ldots (x_k - \text{x-bar}_k)^2$

To get an index of average error variance, divide this quantity by the $df_{wg} = n - k$

This is an estimate of the error variance - variance not accounted for by the IV

Between group variance - what does it estimate?

There will be some between group variance simply due to error variance

If there is no systematic variance, between group variance is an estimate of error variance.

$SSbg = n_1 (\text{x-bar}_1 - GM)^2 + n_2 (\text{x-bar}_2 - GM)^2 + n_k (\text{x-bar}_k - GM)^2$

$df_{bg} = k - 1$

$MS_{bg} = SS_{bg} / k - 1$

So the goal is to determine whether the between group variance, MSbg, is very similar to MSwg, thus making it = to the estimate of the error variance OR whether it is greater than MSwg, thus meaning that the group variable, the IV, is causing additional variance.

Use an F-test to compare the two:

To determine if the F is significant, you look up a critical F in an F-table (Appendix A3). Must know the degrees of freedom and choose an alpha level.

So, the steps are:

1. **Find SSwg, divide by n - k to get MSwg**
2. **Find SSbg, divide by k - 1 to get MSbg**
3. **Calculate F = MSbg / MSwg**
4. **Look up Fcritical**
5. **Compare calculated F to F critical**

Anticipated sizes of F values

Note, if you have two groups, F is = t^2

t's are analogous to 'std. deviations'

so 2 s.ds above the mean is unusual, i.e. t=2 is a bit unusual.

Hence an F of 4 would be considered large

See 1st column of A3, p. 275

Look at table A3 to get an idea of the sizes of F values that are considered significant.

Also, note that the .01 table has higher F values - larger effects are needed if you want to have a lower chance of a Type I error.

The ANOVA summary table

Extending ANOVA to Factorial designs

systematic variance broken up into subparts, for each of the main effects and for the interaction terms

Don't worry about calculation of these

You may see a more extensive ANOVA table, e.g.

How one would talk about the results?

Main effect for A, No main effect for B, Significant interaction.

Followup / Post hoc Tests

When there are more than two levels of an IV, you don't know which was the different one.

post hoc tests or multiple comparison tests

Used AFTER a significant F was found overall.

Between vs. Within-subjects ANOVAs

Just like before. Within-subject designs are more 'powerful' because they minimize error variance

Analysis of these requires special consideration (like the paired-t test rather than the std t-test).

Multivariate ANOVA or MANOVA

Used when more than one DV.

Problem is inflation of Type I error when you have multiple DVs you are analyzing. If 5 DVs you have a higher chance of one of them showing up significant.

Nonexperimental use of ANOVAs

This is fine.

Be aware that just because ANOVA is used it doesn't mean it's a 'true experiment'

Assumptions/Robustness

normal distributions

equal variance between conditions

Single-subject research methods

Why? What's the problem with group research?

Hides data - summarizing across subjects hides individual performance.

The "Average individual" may not be similar to any subject (e.g. 2.1 children).

Note, this does not preclude averaging **within** an individual.

Also, average group behavior can answer some questions about how group on the average will perform - you just can't address individual behavior without further evidence

Examples:

Simulation study of blocking - backward blocking and forward blocking.

Pigeon use of kinematic depth cues

Classic learning curve

Criticisms of group designs

The 'averaging' problem just mentioned

Error variance includes individual differences (intersubject variance) as well as differences within an individual (intra subject variance).

Terms

nomothetic (NA-MA-Thetic)- establish general principles and broad generalizations

idiographic (ID-E-A-Graphic) - seek to describe, analyze and compare individual subject behavior.

Single-subject research

Doesn't imply only a single subject in the study

typically from 3 to 8, but analyzed and discussed as individuals.

Can use statistics (including inferential!!!) as well as graphical techniques.

Typically done under the rubric of Experimental Analysis of Behavior

Journal of the Experimental Analysis of Behavior

Journal of the Applied Analysis of Behavior

Analysis of single-subject data

Graphical analysis

Statistical analysis

EXPERIMENTAL Designs

The search for a STABLE baseline is critical

Also, a SENSITIVE baseline is important

ABA design (also used in other circumstances).

ABC or Multiple-I designs.

Multiple-baseline designs

multiple-baseline across subjects

Multiple-baseline design across behaviors

Multiple-baseline design across situations

Problems and limitations of single-subject designs

- Single-subject data tell us nothing about average performance
- Limit on number of IVs that can be effectively studied within a particular individual
- Limited to application in situations for which a stable and sensitive baseline can be established
- Ethical issues

Ethical Issues in Behavioral Research

Examples in Research

Tuskegee Study

- Begun in 1932 and continued for 40 years
- Purpose was to study the effects of syphilis on untreated subjects
- Participants were 399 black, poor, semiliterate individuals with syphilis

Milgram studies

- 1960s and 1970s
- Participants told that some volunteers would be teachers and others learners in a memory task
- Purpose was not about learning but about coercion

Kammerer's studies on acquired characteristics and their inheritance

- Early 1900s
- He claimed that salamanders not only showed an increasing tendency to have small black spots after being kept on black soil, but that this reduction was being inherited
- His work was doubted, found to be fraudulent (he injected India ink), and he promptly committed suicide

NOTE: Violations are relatively rare.

The Public Health Services for Scientific Integrity examined 118 reported cases

- 27 required full investigations
- 18 involved misconduct
- 3 involved behavioral scientists

Each year, the NSF has found only 2-3 cases of misconduct serious enough to warrant sanctions

- Only 4 behavioral scientists have been cited in recent years
- Steven Breuning, U. of Pittsburgh, 1988
- Arnold Rincover, UNC Greensboro
- Lonnie Mitchell of Coppon State College and Jerusa Wilson

Your basis for moral decisions

Deontological

- you buy into an absolute, universal moral code
- CAREFUL - the universal code is different depending on one's beliefs

Ethical skepticism

morality is relative; decisions are a matter of an individual's conscience

relativism is very common in America and Europe

Utilitarian

cost-benefit analysis

the ends justify the means

Most professional organizations are utilitarian (incl. the APA - American Psychological Association)

NOTE: I find the book's 'questions' too caricatured, esp. of the deontological viewpoint.

Moral dilemmas

Your wife is sick and needs an expensive medicine which you can't afford - she'll die in 3 months if she doesn't get it. The local pharmacy carries it and is not hurting for money. Is it okay to steal the medicine?

NOTE: researchers judge considering the 'higher good' to be a morally superior position.

Human research

If one person must die to find a cure for cancer, would YOU kill them? What if that individual were a criminal? In a coma? What do you think society would conclude?

If tissue from fetuses provide a cure for Alzheimer's, is it acceptable?

Animal research

Animal rights activists take an 'absolute' or deontological perspective on animal research. NOTE: Others can take a deontological perspective and judge this form of research acceptable.

APA takes a utilitarian view but does provide guidelines to minimize suffering and to minimize the number of animals used (*required sample size*).

Basic Ethical Guidelines for Psychological research

APA's Ethical Principles of Psychologists and Code of Conduct (NOTE!!! The book says Code of Content - typo).

Summary of the types of scientific integrity transgressions

Cost-Benefit Analysis: The APA way...

Potential benefits

Basic knowledge

Improvement of research or assessment techniques

Practical outcomes

Benefits for researchers

Benefits for research participants

Potential Costs

Time and effort
Risks to mental or physical welfare
Equipment, salaries, and supplies
Other costs

Balancing Benefits and Costs

The role of the IRB (Institutional Review Board) at one's institution

IRB must have a minimum of 5 members from scientific *and* nonscientific disciplines.

At least 1 member can't be associated with the institution in any other way.

Researchers submit proposals and the IRB evaluates it.

Animal research, revisited

95% of all animal subjects in psych research are rats, mice, and birds

Psychologists' attitudes toward research on animals

Detailed Requirements

Informed Consent

Ensure that subjects know what they're in for
You are not required to divulge everything about the study
INFORMED CONSENT FORM

Problems with obtaining informed consent

Invasion of Privacy

Person has a right to decide "when, where, to whom, and to what extent his or her attitudes, beliefs, and behavior will be revealed"

Public behavior okay
p. 321 of text - go through

Freedom from coercion to participate

Subjects should not be *required* to participate (e.g. as part of a course)

Subjects can drop out at any time during a study.
Milgram's shock study

Deception in Research

Very common
Examples
Concrete example: implicit learning
Objections
HOWEVER - Over 90% of subjects realize deception is sometimes necessary and report positive feelings.

Debriefing
Alternatives to deception

Confidentiality

Identities must be kept secure - no reporting of individual subject information.

Video tapes may need to be destroyed after a certain period of time and they are not to be freely distributed - requires consent of the participant

Common Courtesy

showing up on time, being prepared, behaving warmly, showing appreciation.

So – what do you think of the following? Ethical or not?

Children who exercise regularly achieve more in school than children who don't exercise regularly.

Does making a correlational vs. experimental study matter in your decision?

College students whose parents are divorced declare a major later in their college careers than those whose parents are not divorced.

When someone else is also in a public bathroom, people are more likely to wash their hands than when they are alone.

People who ride bikes six days a week for a half hour lose more weight than those who ride three times a week for an hour.

Again, does correlational vs. experimental matter to your decision?

***** One of the key principles a researcher must follow is to avoid even the APPEARANCE of unethical behavior.**

Research Dissemination

How?

Unpublished manuscript

Conference Poster

Possible Abstract or Proceedings paper

Conference Presentation

also abstract or proceedings possible

Book chapter

Editors "invite" contributors

Book

Journal publication

Minor journal

Major journal

Rejection rates - APA journals, 2000

Evaluating Psychological Research - Consider the source

Where published

Who published

Determining authorship

- **Who gets included and in what order?**
- **Usually, simply doing what you're told doesn't guarantee authorship - must contribute intellectual property**
  - e.g., hired programmers, people who run RPs, secretarial assistants, don't get authorship
  - exception - complex statistical analysis
- **Order of authorship should depend on depth of contribution**
  - Who puts the pen to paper is very important in determining order of authorship
  - Other important contributions include:

Publication Life Cycle

- **Research Completed**
- **Write it up**
- **Submit it to journal editor for review: "submitted"**
  - Blind vs. non-blind reviews
  - Editor determines the reviewers
  - Reviewed by 2-4 individuals
  - Editor compiles Reviews, and makes recommendation
- **If necessary, revise ms. and resubmit**
- **Acceptance! : "in press"**
- **Receive galley proofs for comment (1 day turnaround)**
- **Publication: "(1996)"**

Life cycle a bit different for a Book chapter

- **Editors of the book do all the reviewing**
- **Typically little chance for rejection since you were invited to contribute**

Publication benefits?

- **Book editors and contributors receive minor funds (usu. counted in dollars, not hundreds of dollars).**
- **No remuneration for journal publication**
- **Principal benefit - career advancement**